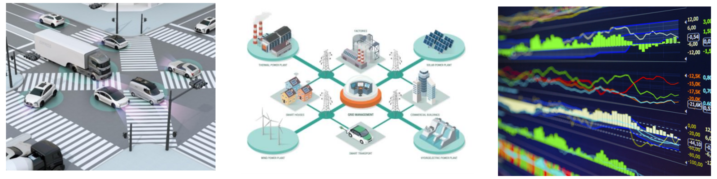
Low Complexity Homeomorphic Projection to Ensure Neural-Network Solution Feasibility for Optimization over (Non-)Convex Set

Enming Liang[1], Minghua Chen[1], Steven H. Low[2]
1. *City University of Hong Kong*, 2. *California Institute of Technology*

Introduction

Real-time constrained optimization is popular

Autonomous driving | Power network | High-freq. trading



Solve real-time optimization is challenging

$$\min_{x\in\mathbb{R}^n} f(x,\theta) \quad \text{s.t. } x \in \mathcal{K}_\theta$$

θ: input parameter
x_θ^*: optimal solution

Iterative strategy

$$x_\theta^{n+1} = x_\theta^n - g(x_\theta^n)$$

- Conventional methods struggle
  - High run-time complexity
  - Not suitable for real-time operations

Input-solution mapping

$$x_\theta^* = F(\theta)$$

- Neural-network scheme is promising
  - Low run-time complexity
  - Not guarantee prediction feasibility

Ensure NN solution feasibility is non-trivial

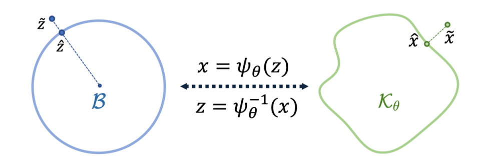
Related work and our contributions

Existing Study	Solution Feasibility Guarantee	Bounded Optimality Loss	Low un-Time Complexity
Penalty approach	✗	✗	✓
Projection approach	✓	✓	✗
Sampling approach	✓	✓	✗
Preventive learning	✓	✗	✓
Gauge mapping	✓	✗	✓
Homeomorphic Projection	✓	✓	✓

HP is the first scheme with feasibility guarantee over ball-homeomorphic sets, bounded optimality loss, and low run-time complexity.

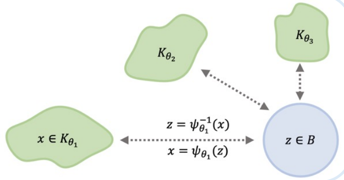
Homeomorphic Projection Framework

Minimum-Distortion Homeomorphic mapping



Projection for feasibility

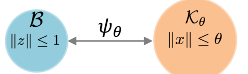
- Over a general set: hard
- Over a ball: easy
- Over ball-homeomorphic sets
  - ✓ any compact convex set
  - ✓ a class of non-convex set



- **MDH mapping**

$$\min_{\psi_\theta\in\mathcal{H}^n} \log \mathrm{D}(\psi_\theta)$$
$$\text{s.t.} \quad \mathcal{K}_\theta = \psi_\theta(\mathcal{B})$$

- **Homeomorphic mapping**: ψ
  - One-to-one mapping
  - Continuous in both directions
- **Distortion**: $\mathrm{D}(\psi) = k_2/k_1 \geq 1$
  - $k_1 = \inf_{z_1,z_2}\{||\psi(z_1) - \psi(z_2)||/||z_1 - z_2||\}$
  - $k_2 = \sup_{z_1,z_2}\{||\psi(z_1) - \psi(z_2)||/||z_1 - z_2||\}$



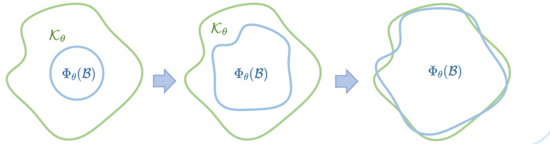
- Low-distortion HM
  - ψ_θ^1: $x = \theta z$
  - $\mathrm{D}(\psi_\theta^1) = 1$
- High-distortion HM
  - ψ_θ^2: $x = \theta \mathrm{R}(||z||)z$
  - $\mathrm{D}(\psi_\theta^2) \approx 2.5$

Unsupervised training with INN for MDH mapping

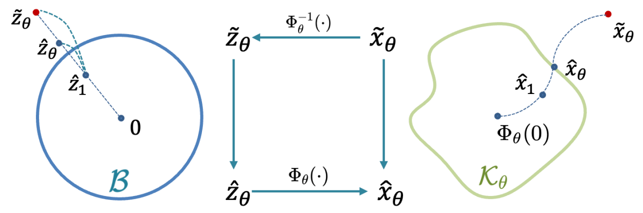
- **Invertible neural network**
  - Composition of many designed invertible layers (e.g., coupling layer)
  - Universal approximator for differentiable homeomorphic mapping
- **Loss function for training INN**
  - One conditional INN learn all θ-dependent MDH mappings

$$\mathcal{L}(\Phi_\theta) = \widehat{\mathrm{V}}(\Phi_\theta(\mathcal{B})) - \lambda_1 \mathrm{P}(\Phi_\theta(\mathcal{B})) - \lambda_2 \widehat{\mathrm{D}}(\Phi_\theta)$$

Volume maximization | Penalty for $\Phi_\theta(\mathcal{B}) \subseteq \mathcal{K}_\theta$ | Distortion minimization



Conduct bisection to recover feasibility



- Step 1: map an infeasible prediction:
  $\tilde{z}_\theta = \Phi_\theta^{-1}(\tilde{x}_\theta)$
- Step 2: bisection for α such that:
  $\alpha^* = \sup_{\alpha\in[0,1]}\{\Phi_\theta(\alpha\cdot\tilde{z}_\theta) \in \mathcal{K}_\theta\}$
- Step 3: find the near boundary point:
  $\hat{x}_\theta = \Phi_\theta(\alpha^*\cdot\tilde{z}_\theta)$

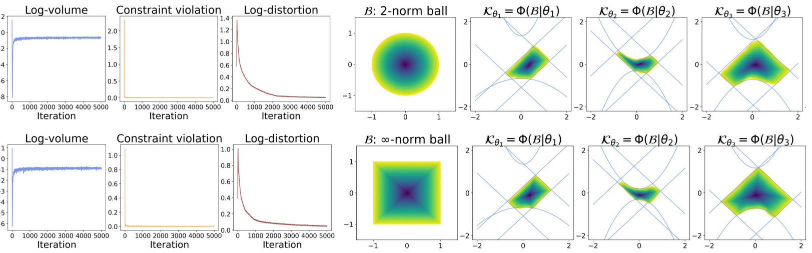
Experiments

Learning MDH mappings

- **Constraint set**

$$\mathcal{K}_\theta = \{x \in \mathbb{R}^2 \mid x^\top Q x + q^\top x + b \leq 0, \quad \theta = [Q,q,b]\}$$

- Training: $\mathcal{L}(\Phi) = \mathbb{E}_\theta[\mathcal{L}(\Phi_\theta)]$
- Testing: $\mathcal{K}_{\theta'} = \Phi(\mathcal{B}|\theta')$



Recover NN solution feasibility

- Optimization problems:
  - QCQP, SDP, and AC-OPF
- Compared methods:
  - Projection, Warm-start, Gradient descent, and H-Projection

	Feasibility: feas. rate %	Feasibility: ineq. vio. 1-norm	Feasibility: eq. vio. 1-norm	Optimality: sol. err. %	Optimality: infeas. sol. err. %	Optimality: obj. err. %	Optimality: infeas. obj. err. %	Speedup: Total ×	Speedup: Post. ×
Convex QCQP: $n=200,\ d=100,\ n_{\text{eq}}=100,\ n_{\text{ineq}}=100$									
NN	54.49	0.163	0	8.16	8.23	3.05	2.96	795657.1	–
NN+WS	100	0	0	4.41	0	1.7	0	2.1	1
NN+Proj	100	0	0	8.15	8.23	3.07	3	2.1	1
NN+D-Proj	56.54	0.023	0	8.15	8.21	3.06	2.98	10.8	4.9
NN+H-Proj	100	0	0	8.36	8.67	3.33	3.58	1618.5	738.8
SDP: $n=15\times15,\ d=100,\ n_{\text{eq}}=100,\ n_{\text{ineq}}=1$									
NN	74.02	11.43	0	6.77	6.99	4.08	3.7	21440.2	–
NN+WS	100	0	0	4.96	0	3.12	0	1.5	0.4
NN+Proj	100	0	0	6.60	6.31	4.43	5.06	1.5	0.4
NN+D-Proj	87.7	5.69	0	6.76	6.94	4.08	3.7	2.6	0.7
NN+H-Proj	100	0	0	7.49	9.76	4.94	7.03	87.6	22.8
118-node AC-OPF: $n=344,\ d=236,\ n_{\text{eq}}=236,\ n_{\text{ineq}}=452$									
NN	73.24	0.006	0	1.27	1.23	0.24	0.23	178.2	–
NN+WS	100	0	0	0.94	0	0.18	0	3.6	1
NN+Proj	100	0	0	1.55	2.31	0.24	0.23	3.8	1
NN+D-Proj	87.79	0.0001	0	1.26	1.23	0.24	0.23	4.9	1.4
NN+H-Proj	100	0	0	1.41	1.78	0.34	0.63	24.6	7.6